

## Novel genetic-based negative correlation learning for estimating soil temperature

S. M. R. Kazemi, Behrouz Minaei Bidgoli, Shahaboddin Shamshirband, Seyed Mehdi Karimi, Mohammad Ali Ghorbani, Kwok-wing Chau & Reza Kazem Pour

**To cite this article:** S. M. R. Kazemi, Behrouz Minaei Bidgoli, Shahaboddin Shamshirband, Seyed Mehdi Karimi, Mohammad Ali Ghorbani, Kwok-wing Chau & Reza Kazem Pour (2018) Novel genetic-based negative correlation learning for estimating soil temperature, Engineering Applications of Computational Fluid Mechanics, 12:1, 506-516, DOI: [10.1080/19942060.2018.1463871](https://doi.org/10.1080/19942060.2018.1463871)

**To link to this article:** <https://doi.org/10.1080/19942060.2018.1463871>



© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 26 Apr 2018.



Submit your article to this journal [↗](#)



Article views: 1334



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 6 View citing articles [↗](#)

# Novel genetic-based negative correlation learning for estimating soil temperature

S. M. R. Kazemi<sup>a</sup>, Behrouz Minaei Bidgoli<sup>b</sup>, Shahaboddin Shamshirband <sup>c,d</sup>, Seyed Mehdi Karimi<sup>e</sup>, Mohammad Ali Ghorbani<sup>f,g</sup>, Kwok-wing Chau<sup>h</sup> and Reza Kazem Pour<sup>f</sup>

<sup>a</sup>Department of Industrial Engineering, Birjand University of Technology, Birjand, Iran; <sup>b</sup>Department of Computer Engineering, Iran University of Science and Technology, Tehran, Iran; <sup>c</sup>Department for Management of Science and Technology Development, Ton Duc Thang University, Ho Chi Minh City, Viet Nam; <sup>d</sup>Faculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City, Viet Nam; <sup>e</sup>Department of Mathematics, Jouybar branch, Islamic Azad University, Jouybar, Iran; <sup>f</sup>Department of Water Engineering, Tabriz University, Tabriz, Iran; <sup>g</sup>Engineering Faculty, Near East University, Mersin, Turkey; <sup>h</sup>Department of Civil and Environmental Engineering, Hong Kong Polytechnic University, Hong Kong, People's Republic of China

## ABSTRACT

A genetic-based neural network ensemble (GNNE) is applied for estimation of daily soil temperatures (DST) at distinct depths. A sequential genetic-based negative correlation learning algorithm (SGNCL) is adopted to train the GNNE parameters. CLMS algorithm is used to achieve the optimum weights of components. Recorded data for two different stations located in Iran are used for the development of the GNNE models. Furthermore, the GNNE predictions are compared with the existing machine-learning models. The results demonstrate that GNNE outperforms other methods for the prediction of DSTs.

## ARTICLE HISTORY

Received 14 December 2017  
Accepted 9 April 2018

## KEYWORDS

Daily soil temperature; neural network ensemble model; negative correlation learning; genetic algorithm; estimation

## 1. Introduction

Accurate estimation of soil temperature plays a vital role in many scientific fields such as agriculture, hydrology, geotechnics, solar and geothermal energy. Soil temperature depends on several factors such as meteorological conditions, physical soil parameters, topographical parameters, and further hydrological parameters of time and depth.

The variation of soil temperature at different depths provides useful information about the land-surface ecosystem processes, environmental and climatic conditions, climate change, and production of crops. Furthermore, soil temperature controls the interactive processes between ground and atmosphere. Although many meteorological variables including relative humidity, air temperature and atmospheric pressure are measured in meteorological sites, measurements of soil temperature data or its spatial variability are not usually available. Consequently, developing theoretical methods to estimate soil temperature from the existing meteorological data is vital.

Soil temperature modeling may be conducted using different approaches: (1) physical methods based on heat transfer mechanism in soil, (2) empirical methods that are commonly regression-based equations between soil

temperature and other available measured meteorological variables, and (3) hybrid methods combining physical and empirical methods. However, the complex inter-relationship between soil temperature and other climatic parameters make accurate prediction of soil temperature a difficult task.

Artificial Intelligence (AI) techniques are considered as well-known tools to deal with the complex problems (Rumelhart & McClelland, 1986). In the last two decades the soft-computing methods have been successfully used in the soil time series predictions such as modeling and prediction of soil salinity (Shahabi, Jafarzadeh, Neyshabouri, Ghorbani, & Valizadeh Kamran, 2017), estimation of field capacity, and permanent wilting point (Ghorbani et al., 2017). (Mihalakakou, 2002) used two models to estimate soil temperature in Dublin and Athens. He reported Artificial Neural Network (ANN) as an efficient tool, yet the analytical model shown to deliver more precision. Regressions and ANN methods were employed by (Bilgili, 2010) to calculate monthly soil temperature at five different depths for Adana, Turkey where the efficiency of ANN models reported higher compared to the regression models. In (Ozturk, Salman, & Koc, 2011), a feed-forward ANN models developed for estimating soil temperature

at different depths in several stations of Turkey. Their outcomes provide a favorable precision for soil temperature estimation. Tabari (Tabari, Sabziparvar, & Ahmadi, 2011) estimated a Daily Soil Temperature (DST) at six depths in IRAN using two regression methods. They demonstrated that ANN is superior to the MLR. In another attempt, the Feedforward Neural Network (FNN) and nonlinear auto regressive neural network with exogenous input (NARX) methods were implemented to forecast weekly soil temperature in Sri Lanka (Napagoda & Tilakaratne, 2013). The NARX model found out to be more precise. (Hosseinzadeh Talaei, 2014) conducted the coactive neuro-fuzzy inference system method to estimate DST in two different region of Iranian stations. It was demonstrated that this method is an adequate method to provide accurate estimation of soil temperature. Recently, (Gill & Singh, 2015) proposed a precise ANN-genetic algorithm to forecast DST.

Chau (2007) developed a high accuracy based on PSO method to train Multilayer Perceptron (MLP) for real-time water level estimation in Shing Mun River of Hong Kong. Three evolutionary algorithms such as DE, ABC and ACO proposed by (Chen, Chau, & Busari, 2015) for downstream river flow forecasting. A ground-water level prediction proposed by (Gholami, Chau, Fadaee, Torkaman, & Ghaffari, 2015) which uses ANN the period from 1912 to 2013. A Binary-coded swarm optimization applied for different gaging station in the northern United states to determine the actual base-flow component expected for most of the considered gages (Taormina, Chau, & Sivakumar, 2015). (Wang, Xu, Chau, & Lei, 2014) proposed a fuzzy-based approach to evaluate the quality of river water. A genetic algorithm (GA)-based on ANN employed for flood forecasting in Yangtze River, China (Wu & Chau, 2006). (Uyumaz, Danandeh Mehr, Kahya, & Erdem, 2014) researched on linear genetic programming (LGP) technique to estimate rectangular side weirs in circular channel. Furthermore, (Bonakdari, Baghalian, Nazari, & Fazli, 2011) studied on numerical analysis for prediction of flow field in a 90 bend.

A survey of the articles reveals the appealing application of the AI methods in soil temperature prediction. Possibly, an efficient approach to solve such complex real-world problems is to create a group of predictors by combining the AI prediction models in order to enhance the overall performance of prediction system (Masoudnia, Ebrahimpour, & Arani, 2012). Joint learning methods mainly aim at simplifying a prediction task through a convergence prediction outcome for the base data (García-Pedrajas, Maudes-Raedo, García-Osorio, & Rodríguez-Díez, 2012). In this context, neural network ensemble (NNE) is one of the dominant

ensemble methods (Hansen & Salamon, 1990). NNE involves a certain number of neural networks (Tian, Li, Chen, & Kou, 2012; Zhai, Xu, & Wang, 2012). Usually, these methods have two separate phases of creating NN components and combining those components. Experimental and theoretical research revealed that the NNE model is more useful if estimations of model components have a negative correlation. In this case there is no relation between its components and NNE performs less effective. In fact the algorithm will act with no efficiency if there is positive correlation between the components (Brown, Wyatt, Harris, & Yao, 2005). A common method to create negatively correlated and accurate components, appropriate for the NNE model is penalty method through manipulation of the training dataset. Bootstrap aggregating are common methods of building ensembles. These methods manipulate the examples to train individual NN components sequentially and independently (Breiman, 1996; Freund & Schapire, 1996).

Negative correlation learning (NCL) (Liu & Yao, 1999) is a common consequence method to develop diverse components which boost a correlation penalty terms for component of cost functions to reduce mean square error (MSE) of every component and the error correlation of ensemble (Masoudnia et al., 2012). There is a regularization term in NCL that helps balancing the bias-var-cov trade-off and enhancing the ability to generalize. The typical NCL has certain constants. A study showed that NCL only drives away the ensemble components from the mean and not certainly from one to another (McKay & Abbass, 2001). Consequently, the NCL's low variety is not unexpected. Moreover, the back-propagation algorithm (BP) can be employed to train the neural network components in the NCL method. The BP algorithm is often stuck in a local minimum and it is unable to find a global minimum (Asadi, Hadavandi, Mehmanpazir, & Nakhostin, 2012).

The classic NCL static combiner approaches are not able to model of the local fitness of its components. Lee (Lee, Kim, & Pedrycz, 2012) introduced a novel selective ensemble of NN-Negative Correlation. A series of component networks utilized to minimize the generalization error and increase the negative correlation. Another NCL model was proposed by (Alhamdoosh & Wang, 2014) to create ensembles capable of sound generalization by regulating the incongruity between outputs of the base learners. They used the random vector functional link networks (RVFL) to find a better solution.

In this study, a new ANN ensemble method entitled as genetic-based NNE (GNNE) is proposed by combining an NCL model (Hadavandi, Shahrabi, & Hayashi, 2015) and evolutionary algorithm to calculate daily soil temperature at 5, 10, 20, 30, 50 and 100 cm depths. A

sequential genetic-based NCL (SGNCL) (Kazemi, Hada-vandi, Shamshirband, & Asadi, 2016) algorithm is introduced to train the components of the GNNE. SGNCL uses genetic algorithm (GA) as a global search method instead of the BP algorithm to tune the weights of network. A CLMS algorithm considered to find best components weights in combination module of the GNNE. To build the soil temperature prediction models, a daily weather dataset of two Iranian stations with different climate characteristics are utilized. The performance of the GNNE model is compared with our previous work which used an extreme learning machine (ELM) and hybrid self-adaptive evolutionary (SaE)-ELM (SaE-ELM) (Nahvi, Habibi, Mohammadi, Shamshirband, & Al Razgan, 2016).

## 2. Methodology

In this section, the implemented technique and related formulation process for estimating soil temperatures are described.

### 2.1. The suggested genetic-based neural network ensemble

To comprehend the possibility of generalizing the ensemble, Theorem 1 from (Naonori Ueda & Nakano, 1996) is recalled and expressed below:

**Theorem 2.1.** *The subsequent bias-variance-covariance decomposition is employed to express an ensemble's generalization error averaged over total potential  $D_N$  (training datasets) of size  $N$ , where the same dataset trains all components:*

$$E_{f_{ens}}(D_v) = \frac{1}{N} \sum_{n=1}^N \left[ \frac{1}{T} \overline{Var}(X_n) + \frac{T-1}{T} \overline{Cov}(X_n) + \overline{Bias}(X_n)^2 \right] \quad (1)$$

Here,  $D_v$  represents the validation dataset. Eqs. (2)–(4) are used to determine  $\overline{Cov}(X_n)$ ,  $\overline{Bias}(X_n)$  and  $\overline{Var}(X_n)$  as below:

$$\overline{Var}(X_n) = \frac{1}{T} \sum_{t=1}^T E_{D_N} \{ (f_t(X_n) - E_{D_N} \{ f_t(X_n) \})^2 \} \quad (2)$$

$$\overline{Cov}(X_n) = \frac{1}{T(T-1)} \sum_{t=1}^T \sum_{k \neq t} E_{D_N} \{ (f_t(X_n) - E_{D_N} \{ f_t(X_n) \}) (f_k(X_n) - E_{D_N} \{ f_k(X_n) \}) \} \quad (3)$$

$$\overline{Bias}(X_n) = \frac{1}{T} \sum_{t=1}^T (E_{D_N} \{ f_t(X_n) \} - y(X_n)) \quad (4)$$

Here,  $N$  represents the number of training cases, the ensemble size is shown by  $T$ ,  $f_t(X_n)$  is the output of component  $t$  and  $f_{ens}(X_n) = (1/T) \sum_{t=1}^T f_t(X_n)$  is the ensemble model for  $X_n$ . (Geman, Bienenstock, & Doursat, 1992) provided the proof for Theorem 1. Based on Eq. (1), the disagreement between outputs of ensemble components is controlled by explicit handling the covariance term  $\overline{Cov}(X_n)$ , thus developing a better generalized ensemble model. Eqs. (5) and (6) show the NCL's cost functions for  $t$  ( $1 \leq t \leq T$ ):

$$E_t = \left[ \sum_{n=1}^N (f_t(X_n) - y(X_n))^2 \right] + \lambda P_t \quad (5)$$

$$P_t = \sum_{n=1}^N (f_t(X_n) - f_{ens}(X_n)) \sum_{L=1, L \neq t}^T (f_L(X_n) - f_{ens}(X_n)) \\ = - \sum_{n=1}^N (f_t(X_n) - f_{ens}(X_n))^2 \quad (6)$$

$\lambda$  regulates the trade-off between the penalty function and objective, and  $P_t$  is the coefficient of error correlation penalty. The ensemble components can be trained individually using the BP algorithm when  $\lambda = 0$ . Therefore, we propose the SGNCL algorithm to train the GNNE model components. It has key characteristics as below:

- 1) In the error function of component  $t$ , the GNNE uses a moving average for the correlation penalty term that calculates the error correlation between previous components and itself ( $1, 2, \dots, t-1$ ), and successively locates every component in a position relative to preceding components, hoping that GNNE components trained by the SGNCL move away from each other, hence maximizing their diversity. Eqs. (7) and (8) express the suggested cost functions in the SGNCL for training  $t$ :

$$E_t = \sum_{k=1}^P \sum_{n=1}^N (f_{tk}(X_n) - y_k(X_n))^2 + \lambda P_t \quad (7)$$

$$P_t = \sum_{k=1}^P \left[ \sum_{n=1}^N (f_{tk}(X_n) - y_k(X_n)) \sum_{L=1}^{t-1} (f_{Lk}(X_n) - y_k(X_n)) \right] \quad (8)$$

In these equations,  $y_k(X_n)$  and  $f_{ik}(X_n)$  are the actual and predicted values of  $k$  target of instance  $n$ , respectively ( $1 \leq k \leq P$ ), and the quantity of training instances is represented by  $N$ .

- 2) The GNE employs (GA) to tune the weights of the NN experts (Asadi et al., 2012). Thus, we used GA in the suggested SGNCL algorithm in this paper to obtain the best set of network in the subsequent phases:

#### Phase 1 – Encoding

Every gene signifies the weight between two neurons in various layers. As shown in Figure 1, a chromosome has a series of genes. As such, for a normal feed-forward neural network one neuron in output layer, with two neurons in hidden layer, and three neurons in input layer, the 1st gene is the weight between neuron 4 and 1, namely  $W_{14}$ . The 2nd gene is the weight between neuron 5 and 1, namely  $W_{15}$  and so on. Real-number form is used to signify the connection weights. So in this way we encoded the weights of feed-forward neural network by each gene in chromosome as it seen below.

#### Phase 2 – Create the primary population

The primary population ( $N_{pop}$ ) is created arbitrarily. Every single initial weight is developed randomly between  $-1$  and  $1$ . Hence there are first population which created discretionary and first weight which distributed uniformly between  $-1$  and  $1$ .

#### Phase 3 – Obtaining the fitness values

These values are calculated based on the  $E_t$  in this case the fitness values obtained and they established on  $E_t$ .

#### Phase 4 – Selection mechanism

The truncation selection scheme was used. In this mechanism, individuals are organized based on their fitness to choose the best individuals for parents. The population's quantity that is designated as parents is

called the truncation threshold. Next, binary tournament selection scheme is adopted to choose parents for creating new off springs through sue genetic operators. Two members of the population are chosen in binary tournament selection randomly, then the fitness function is associated to select the best for one parent according to the fitness value. Moreover, other parents are chosen through a similar procedure. Selecting randomly parents and comparing their fitness according to fitness value which is depended on  $E_t$  as well.

#### Phase 5 – Genetic operators

The binary tournament selection is applied to the top-half best-performing persons to choose individuals for crossover. One-point crossover is done on the designated parents to create offspring. Next, the newly produced offspring replace individuals with poor performance. Likewise, uniform mutation is employed, which substitutes a randomly chosen gene with a value selected from a uniform random distribution between lower and upper domain limits for the gene. The probability of mutation must, therefore, be small enough to avoid degenerate of the performance of GA.

#### Phase 6 – Replacement

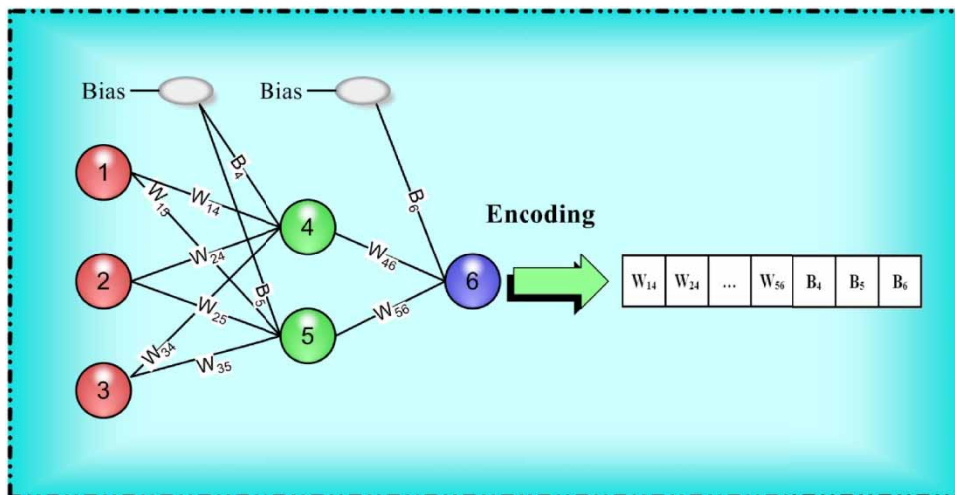
The newly generated offspring replace the current population to create the succeeding generation. Due to this replacement new generations produced and it occur until reaching the stop phase.

#### Phase 7 – Stopping measures

In case the quantity of generations exactly resembles the maximum number of generation, then stop; or else move to Phase 3.

## 2.2. Combination method

Following training of the MLP components, a combination method is needed to combine the outputs. This



**Figure 1.** Coding a neural network component as a chromosome.

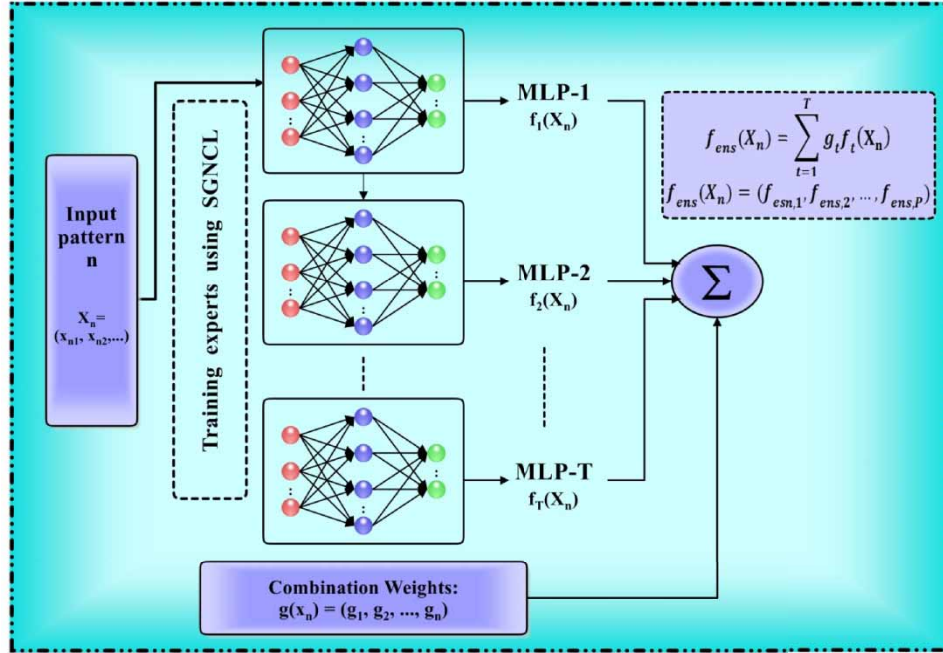


phase aims to fully take advantage of diversity's positive effects and evade from the negative outcomes (Yang, Zeng, Zhong, & Wu, 2013). Similar to ensemble of MLP components, weighted and simple averaging are the most popular combination methods for regression problems. Nevertheless, it is worth exploring methods of finding suitable weights for component of NNE (Tian et al., 2012; Yang et al., 2013). Eq. (9) computes the last GNNE model

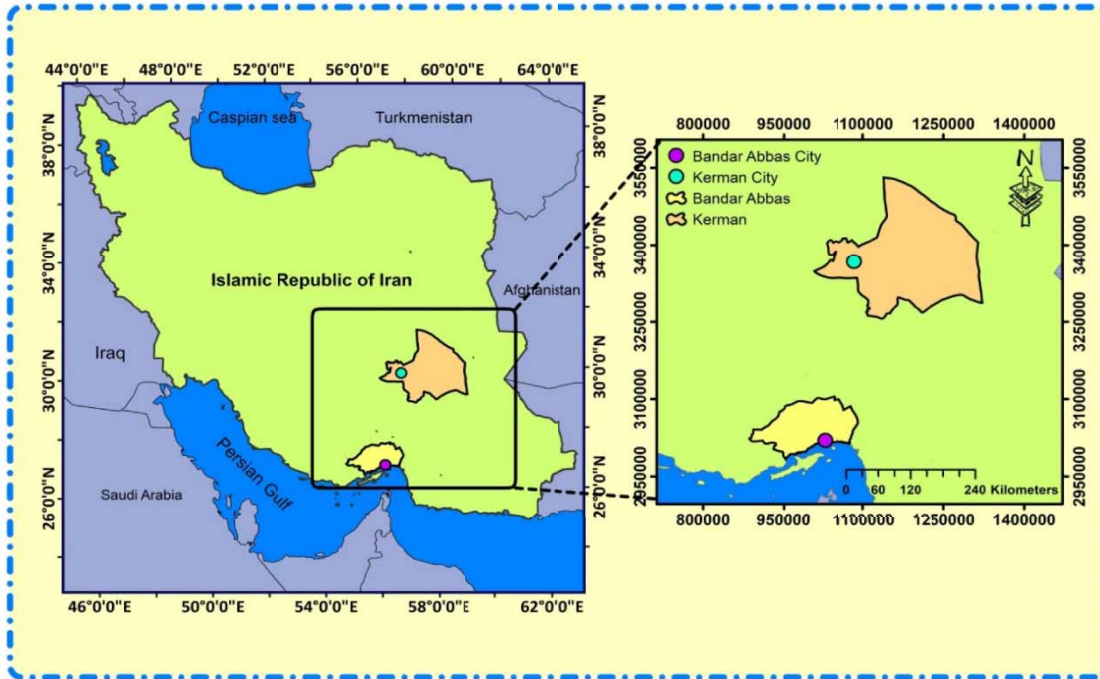
prediction with  $T$  components for target  $k$  ( $1 \leq k \leq P$ ):

$$f_{ens,k}(X_n) = \sum_{t=1}^T g_{tk} f_{tk}(X_n) \quad (9)$$

Here, the ensemble model predicts  $f_{ens,k}(X_n)$  and  $f_{tk}(X_n)$  values and  $t$  for  $k$  target of  $n$  instance, and  $g_{tk}$  represents the weight of  $t$  for  $k$ .



**Figure 2.** Diagram of the suggested GNNE model with  $T$  components in the testing step.



**Figure 3.** The geographic location of Bandar Abbas and Kerman city.

In an NNE model, there are several ways to measure the weight parameters of components. Two main categories exist. One includes methods that use the CLMS with a mathematical outline to find optimal weights. The combination weights are obtained through an optimization technique, such as the Lagrange multipliers (Ueda, 2000). In the second category, an evolutionary algorithm is used to discover an optimal combination weights based on the outcomes of every single component of the training dataset (Nabavi-Kerizi, Abadi, & Kabir, 2010). The CLMS was used in this work to get the optimal components' weights (Hadavandi, Shahrabi, & Shamshirband, 2015). The  $g_{tk}$  weights are calculated so that to lessen the probable quadratic deviation of  $E$  for the assumed training set. Calculation of weight parameters for  $k$  target feature ( $1 \leq k \leq P$ ) is expressed in this paper as follows:

$$\text{Minimize } E = \sum_{n=1}^N \left( y_k(X_n) - \left( \sum_{t=1}^T g_{tk} f_{tk}(X_n) \right) \right)^2$$

$$\sum_{t=1}^T g_{tk} = 1 \quad g_{tk} \geq 0 \quad (10)$$

Figure 2 shows the suggested GNNE model with its components.

### 3. Dataset

In this study, the datasets from Bandar Abbas and Kerman cities in Iran were utilized. Bandar Abbas is the capital of Hormozgan province situated at the northern coast of Persian Gulf at 27°13'/N and 56°22'/E, and 10 m

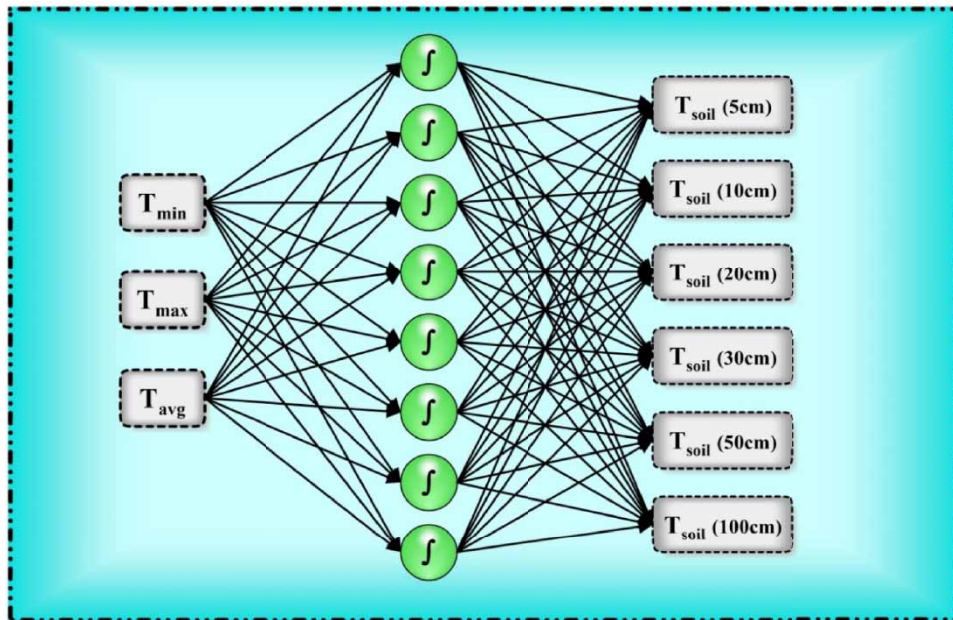
above the sea level. In summer, the highest air temperature is up to 49 °C whereas in winters the minimum air temperatures can reach to 5 °C. Long-term yearly averaged air temperature and relative humidity are 27 °C and 65%. Also, the annual rainfall is approximately 170 mm. In summer, it has one of the maximum average dew points of any location in the world. Bandar Abbas has a hot desert climate categorized as BWh based on Köppen climate classification (Kottek, Grieser, Beck, Rudolf, & Rubel, 2006).

**Table 1.** Obtained statistical parameters for the GNNE model for Bandar Abbas.

Soil Depth (cm)	Training data			Testing data		
	MABE (°C)	RMSE (°C)	$R^2$	MABE (°C)	RMSE (°C)	$R^2$
5	0.840	1.020	0.985	0.850	1.050	0.983
10	0.800	1.010	0.979	0.840	1.040	0.979
20	0.800	1.130	0.988	0.810	1.100	0.981
30	0.920	1.130	0.971	0.970	1.210	0.969
50	1.010	1.360	0.941	1.030	1.320	0.938
100	1.380	1.770	0.84	1.400	1.830	0.835

**Table 2.** Obtained statistical parameters for the GNNE model for Kerman.

Soil Depth (cm)	Training data			Testing data		
	MABE (°C)	RMSE (°C)	$R^2$	MABE (°C)	RMSE (°C)	$R^2$
5	1.430	1.980	0.971	1.560	2.070	0.969
10	1.460	1.910	0.972	1.500	2.001	0.968
20	1.600	2.070	0.961	1.600	2.080	0.954
30	1.700	2.180	0.929	1.710	2.170	0.928
50	1.910	2.400	0.890	1.950	2.430	0.899
100	2.200	2.810	0.770	2.230	2.850	0.779



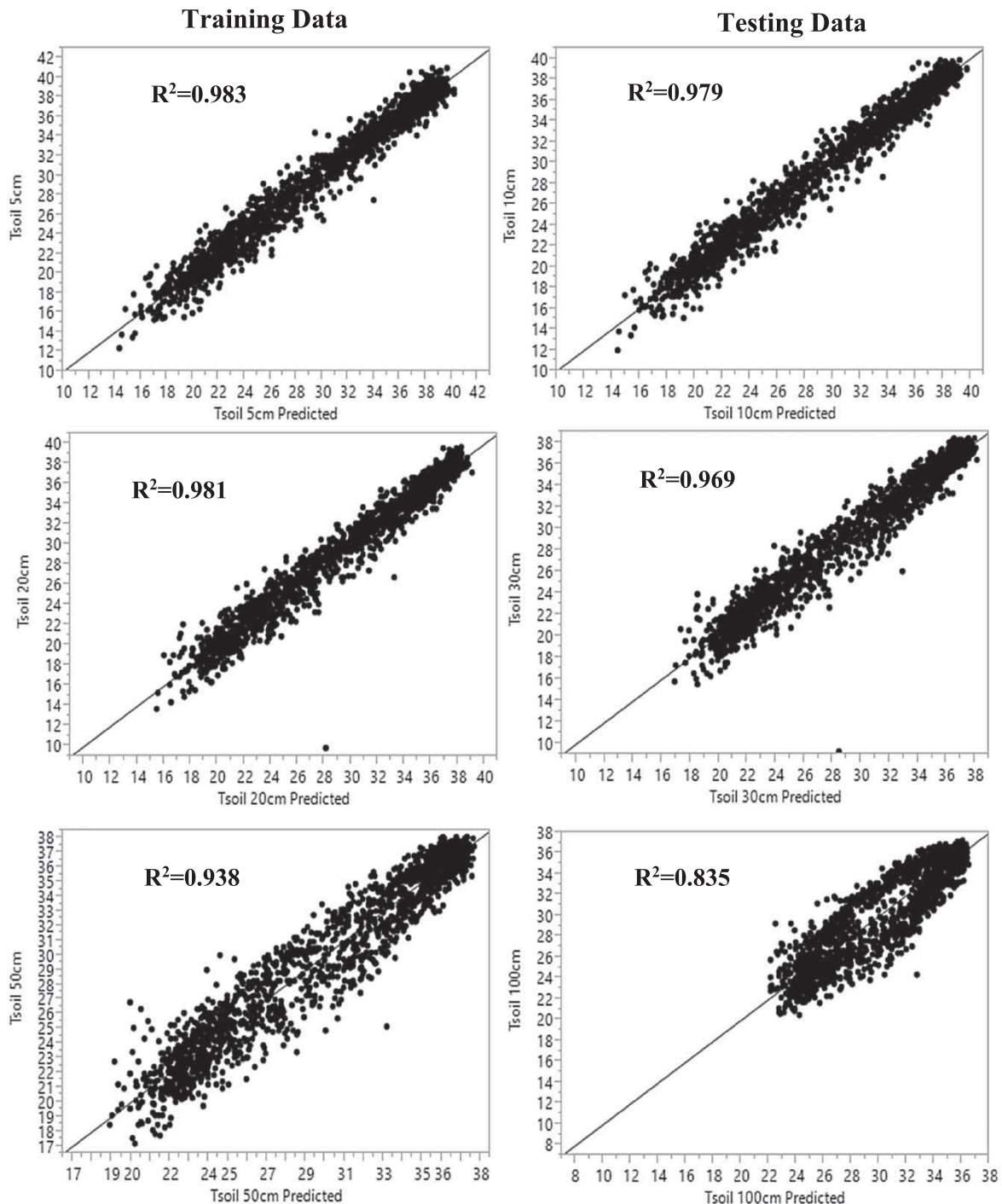
**Figure 4.** The architecture of the components in the GNNE model.

City of Kerman is situated southeast of country at 30°29' N and 57°06' E, and its elevation is 1754 m above the sea level. Its climate condition is generally moderate and dry which experience hot summers because of its placement in the loot desert. Long-term yearly averaged air temperature and relative humidity in Kerman are 15.8 °C and 32%, respectively. Also, the annual average of rainfall is around 140 mm. Kerman has a cold desert

climate classified as BWk on the basis of Köppen climate classification.

In Figure 3, we show the geographic location of the tested cities

For Bandar Abbas, 10 years recorded data from Jan 1996 to Dec 2005, and for Kerman, 7 years recorded data from Jan 1998 to 2004 were used. The daily data collections consist of daily soil temperature (ST) at several

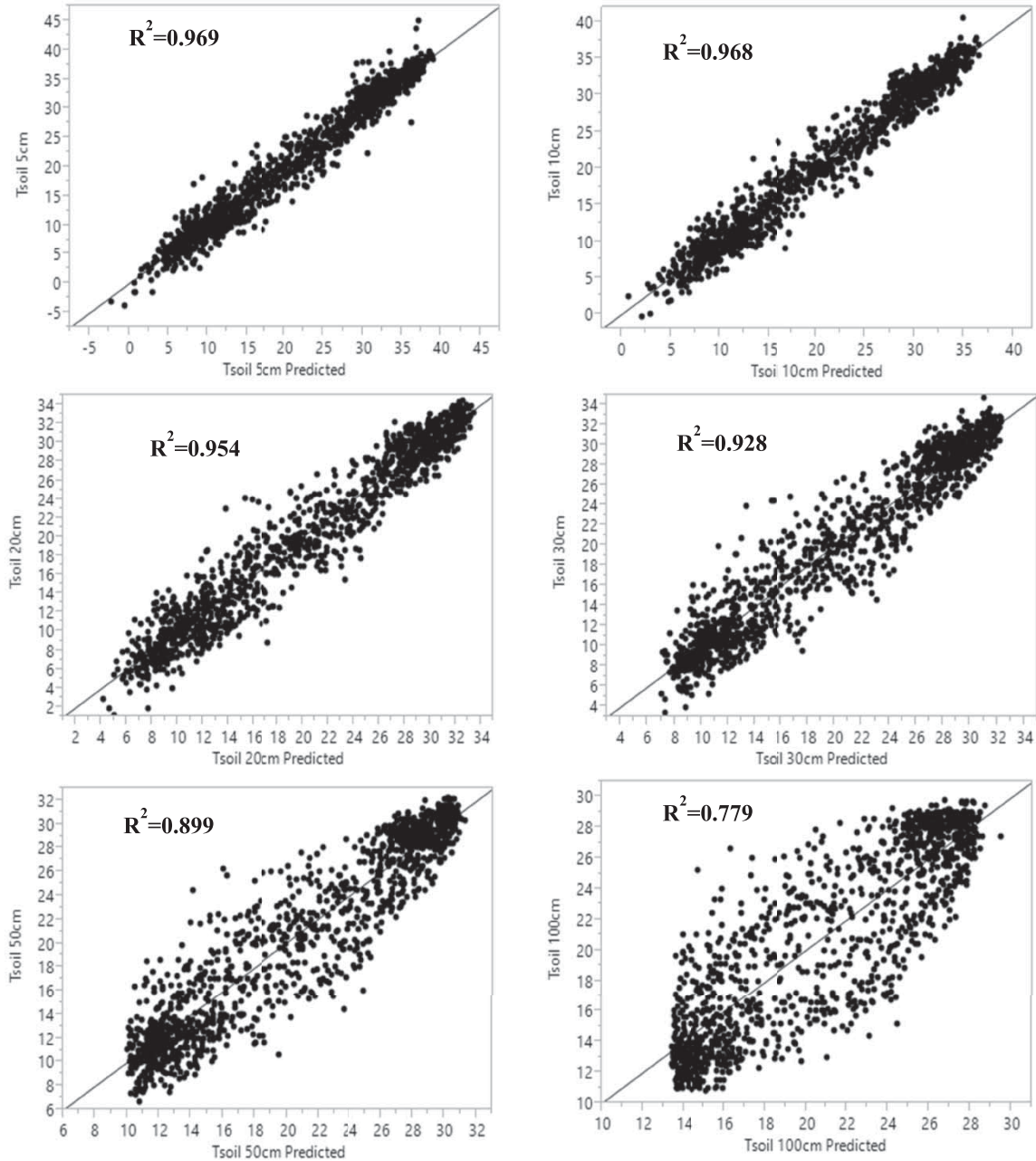


**Figure 5.** Comparison of the estimated daily ST by the GNNE model against measured data for Bandar Abbas.



depths as well as maximum, minimum and average air temperatures. We divided the data collections to two subsets for training (50% of total data) and for testing (50% of total data). The datasets used in this study are part of the data sets utilized in a previous study on estimating daily soil temperatures at Bandar Abbas and Kerman using the SaE-ELM and ELM models (Nahvi et al., 2016). Using self-adaptive evolutionary algorithm to improve the performance of an extreme learning machine for estimating soil temperature. This enables us to validate the performance of the GNNE against other robust AI methods.

To determine how much air temperatures are significant for predicting ST, the correlation coefficient between air temperatures and ST at all depths were calculated. For both stations, air temperatures showed favorable correlations with ST at all depths. It was found that the correlation coefficient for both stations decrease slightly with increase of the soil depth from 5 to 100 cm. For Bandar Abbas, the correlation coefficients of  $T_{\min}$ ,  $T_{\max}$  and  $T_{\text{avg}}$  with ST at different depths were between 0.8602 and 0.9574, 0.8352 and 0.9307 as well as 0.8793 and 0.9793, respectively. In addition, for Kerman the correlation coefficients of  $T_{\min}$ ,  $T_{\max}$  and  $T_{\text{avg}}$  with ST at



**Figure 6.** Comparison of the estimated daily ST by the GNNE model against measured data for Kerman.

different depths were between 0.8280 and 0.9038, 0.8492 and 0.9559 as well as 0.8366 and 0.9722, respectively. These high correlations of air temperatures with ST at different depths indicates that air temperatures are highly influential and suitable for predicting soil temperature. This can be attributed to the fact that air and soil temperatures are determined according to the energy balance at the ground surface.

#### 4. Results and discussion

A series of analyses were carried out by developing the GNNE models and comparing their performance with other models. Multi-Layer perceptron adopts as a NN components of the suggested model. In the hidden layer, the number of nodes was set at 8. There are three inputs in input layer for each modeling and at last outputs that are shown in Figure 4. The hyperbolic tangent function was the activation function employed in the output and hidden nodes and gating network. Figure 4 depicts the architecture of the GNNE model's MLP components.

The performance of the proposed GNNE model is assessed by mean absolute bias error (MABE), root mean square error (RMSE) and coefficient of determination ( $R^2$ ). The model with smaller RMSE and MABE brings more accuracy, whereas bigger  $R^2$  indicates that the model offers more precision. Tables 1 and 2 show the obtained statistical indicators for the GNNE model for Bandar Abbas and Kerman, respectively. The statistical results reveal that for both cases studies, the GNNE model has a favorable performance for estimating daily soil temperature at all depths. The predictions of ST at all depths have very good agreements with the measured data. RMSE and MABE increase and  $R^2$  decreases with increase of the soil depth. This is clearer when the soil depth rises from 50 to 100 cm. The reason is that influence of the considered meteorological variables on ST decrease at higher depths.

Moreover, the estimated daily ST by the GNNE model at all depths are plotted against the obtained data for the testing stage for Kerman and Bandar Abbas (Figure 5 and Figure 6). The plots indicate that there is a good correlation between the estimated and measured data. However, the higher scattering of the data points at higher soil depths implies a decrease of the prediction accuracy with increasing the depth.

In order to have an idea about the prediction performance of GNNE, its accuracy is compared with the classical ELM and combined SaE-ELM proposed in authors' previous study for the same case studies (Nahvi et al., 2016). To fulfill this validation phase, the average values of MABE and RMSE for all six depths achieved for Bandar Abbas and Kerman cities are calculated. Figures

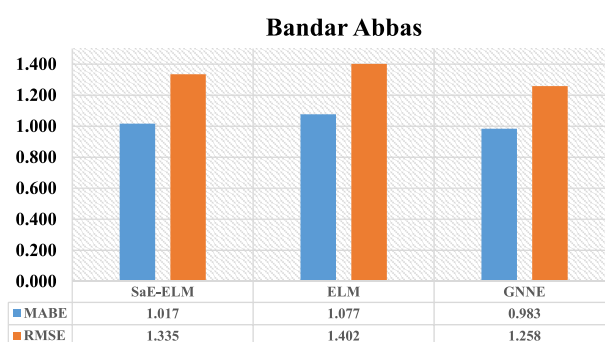


Figure 7. Comparison of the prediction performance of the GNNE, ELM and SaE-ELM models for Bandar Abbas.

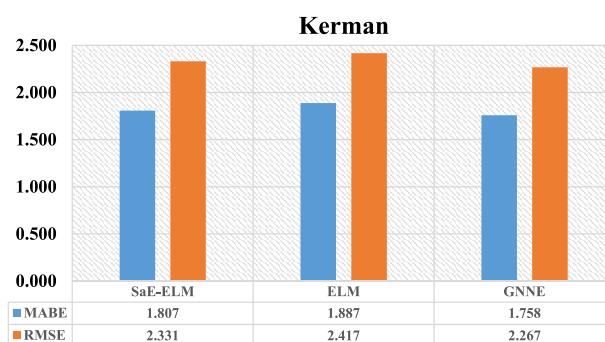


Figure 8. Comparison of the prediction performance of the GNNE, ELM and SaE-ELM models for Kerman.

7 and 8 shows the comparative study results. As seen, the GNNE models notably outperform the SaE-ELM and ELM models on the data for both Bandar Abbas and Kerman cities.

#### 5. Conclusions

A forecast model for soil temperature (especially on daily horizon) is an important decision-support tool in soil engineering, agriculture, meteorology, and geotechnics. Thus, determination of soil temperature at various depths is critical in many research fields. The present research introduces a new GNNE method by coupling NCL with evolutionary algorithms to predict soil temperature at several depths. The proposed method combines different NN prediction models and builds a predictor ensemble. The SGNCL algorithm is implemented for training the components in GNNE, and LMS algorithm integrated for reaching the components' optimal weights. Daily weather datasets of two stations in different regions of Iran with variety of climate conditions are used for developing the models. The proficiency of the GNNE models is assessed by comparing with two evolutionary models such as ELM and SaE-EL. Under this circumstance the statistical evaluations are RMSE and MABE. The results demonstrate

that the GNNE model provides significant level of precision in predicting daily soil temperatures in all depths for both stations as well. For the Bandar Abbas station, the MABE and RMSE values for all depths lie between 0.85 and 1.40 °C, and between 1.05 and 1.83 °C, respectively. The recorded RMSE and MABE and for the Kerman station are ranged between 1.56–2.23 °C and 2.07–2.85 °C, respectively. The results show that the predictions which produced by GNNE models are superior to those made by the SaE-ELM and ELM methods in both stations considering to evaluation criteria. These findings suggest the power of application of GNNE method in predicting daily soil temperature in different depths and it makes a way to encourage researchers for utilizing GNNE model in this scientific area. Future research in this field can focus on the inclusion of other fuzzy systems or neural networks with the NCL process. Making effective data decomposition algorithms to construct diverse components in the GNNE can also introduce another field for the research on GNNE model in further research.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## ORCID

Shahaboddin Shamshirband  <http://orcid.org/0000-0002-6605-498X>

## References

- Alhamdoosh, M., & Wang, D. (2014). Fast decorrelated neural network ensembles with random weights. *Information Sciences*, 264, 104–117. doi:10.1016/j.ins.2013.12.016
- Asadi, S., Hadavandi, E., Mehmanpazir, F., & Nakhostin, M. M. (2012). Hybridization of evolutionary levenberg–marquardt neural networks and data pre-processing for stock market prediction. *Knowledge-Based Systems*, 35, 245–258.
- Bilgili, M. (2010). Prediction of soil temperature using regression and artificial neural network models. *Meteorology and Atmospheric Physics*, 110, 59–70.
- Bonakdari, H., Baghalian, S., Nazari, F., & Fazli, M. (2011). Numerical analysis and prediction of the velocity field in curved open channel using artificial neural network and genetic algorithm. *Engineering Applications of Computational Fluid Mechanics*, 5(3), 384–396.
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24, 123–140.
- Brown, G., Wyatt, J., Harris, R., & Yao, X. (2005). Diversity creation methods: A survey and categorisation. *Journal of Information Fusion*, 6(1), 5–20.
- Chau, K. (2007). A split-step particle swarm optimization algorithm in river stage forecasting. *Journal of Hydrology*, 346(3–4), 131–135.
- Chen, X. Y., Chau, K. W., & Busari, A. O. (2015). A comparative study of population-based optimization algorithms for downstream river flow forecasting by a hybrid neural network model. *Engineering Applications of Artificial Intelligence*, 46, 258–268.
- Freund, Y., & Schapire, R. E. (1996). *Experiments with a new boosting algorithm. The thirteenth international conference on machine learning – ICML*.
- García-Pedrajas, N., Maudes-Raedo, J., García-Osorio, C., & Rodríguez-Díez, J. J. (2012). Supervised subspace projections for constructing ensembles of classifiers. *Information Sciences*, 193, 1–21.
- Geman, S., Bienenstock, E., & Doursat, R. (1992). Neural networks and the bias/variance dilemma. *Neural Computation*, 4(1), 1–58.
- Gholami, V., Chau, K. W., Fadaee, F., Torkaman, J., & Ghaffari, A. (2015). Modeling of groundwater level fluctuations using dendrochronology in alluvial aquifers. *Journal of Hydrology*, 529, 1060–1069.
- Ghorbani, M. A., Shamshirband, S., Zare Haghi, D., Azani, A., Bonakdari, H., & Ebtehaj, I. (2017). Application of firefly algorithm-based support vector machines for prediction of field capacity and permanent wilting point. *Soil and Tillage Research*, 172, 32–38.
- Gill, J., & Singh, S. (2015). An efficient neural networks based genetic algorithm model for soil temperature prediction. *International Journal of Emerging Technologies in Engineering Research (IJETER)*, 3, 1–5.
- Hadavandi, E., Shahrabi, J., & Hayashi, Y. (2015). SPMoe: A novel subspace-projected mixture of experts model for multi-target regression problems. *Soft Computing*, 20(5), 2047–2065. doi:10.1007/s00500-015-1623-7
- Hadavandi, E., Shahrabi, J., & Shamshirband, S. (2015). A novel boosted-neural network ensemble for modeling multi-target regression problems. *Engineering Applications of Artificial Intelligence*, 45, 204–219. doi:10.1016/j.engappai.2015.06.022
- Hansen, L. K., & Salamon, P. (1990). Neural network ensembles. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(10), 993–1001.
- Hosseinzadeh Talaei, P. (2014). Daily soil temperature modeling using neuro-fuzzy approach. *Theoretical and Applied Climatology*, 118(3), 481–489.
- Kazemi, S. M. R., Hadavandi, E., Shamshirband, S., & Asadi, S. (2016). A novel evolutionary-negative correlated mixture of experts model in tourism demand estimation. *Computers in Human Behavior*, 64, 641–655.
- Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, 15(3), 259–263.
- Lee, H., Kim, E., & Pedrycz, W. (2012). A new selective neural network ensemble with negative correlation. *Applied Intelligence*, 37(4), 488–498. doi:10.1007/s10489-012-0342-3
- Liu, Y., & Yao, X. (1999). Ensemble learning via negative correlation. *Neural Networks*, 12(10), 1399–1404. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12662623>
- Masoudnia, S., Ebrahimpour, R., & Arani, S. A. A. A. (2012). Combining features of negative correlation learning with mixture of experts in proposed ensemble methods. *Applied Soft Computing*, 12, 3539–3551.
- McKay, R., & Abbass, H. A. (2001). Anti-correlation: A diversity promoting mechanism in ensemble learning. *The Australian Journal of Intelligence Information Processing Systems*, 7(3), 139–149.
- Mihalakakou, G. (2002). On estimating soil surface temperature profiles. *Energy and Buildings*, 34, 251–259.

- Nabavi-Kerizi, S. H., Abadi, M., & Kabir, E. (2010). A PSO-based weighting method for linear combination of neural networks. *Computers and Electrical Engineering*, 36, 886–894.
- Nahvi, B., Habibi, J., Mohammadi, K., Shamshirband, S., & Al Razgan, O. S. (2016). Using self-adaptive evolutionary algorithm to improve the performance of an extreme learning machine for estimating soil temperature. *Computers and Electronics in Agriculture*, 124(Supplement C), 150–160. doi:10.1016/j.compag.2016.03.025
- Napagoda, N. A. D. N., & Tilakaratne, C. D. (2013). Artificial neural network approach for modeling of soil temperature: A case study for Bathalagoda Area. *Sri Lankan Journal of Applied Statistics*, 13, 39–59.
- Ozturk, M., Salman, O., & Koc, M. (2011). Artificial neural network model for estimating the soil temperature. *Canadian Journal of Soil Science*, 91, 551–562.
- Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the micro structure of cognition, foundations*. Cambridge, MA: MIT Press.
- Shahabi, M., Jafarzadeh, A. A., Neyshabouri, M. R., Ghorbani, M. A., & Valizadeh Kamran, K. (2017). Spatial modeling of soil salinity using multiple linear regression, ordinary kriging and artificial neural network methods. *Archives of Agronomy and Soil Science*, 63(2), 151–160.
- Tabari, H., Sabziparvar, A.-A., & Ahmadi, M. (2011). Comparison of artificial neural network and multivariate linear regression methods for estimation of daily soil temperature in an arid region. *Meteorology and Atmospheric Physics*, 110(3–4), 135–142.
- Taormina, R., Chau, K.-W., & Sivakumar, B. (2015). Neural network river forecasting through baseflow separation and binary-coded swarm optimization. *Journal of Hydrology*, 529, 1788–1797.
- Tian, J., Li, M., Chen, F., & Kou, J. (2012). Coevolutionary learning of neural network ensemble for complex classification tasks. *Pattern Recognition*, 45, 1373–1385.
- Ueda, N. (2000). Optimal linear combination of neural networks for improving classification performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(2), 207–215.
- Ueda, N., & Nakano, R. (1996). *Generalization error of ensemble estimators*. *IEEE international conference on neural networks*, 1996.
- Uyumaz, A., Danandeh Mehr, A., Kahya, E., & Erdem, H. (2014). Rectangular side weirs discharge coefficient estimation in circular channels using linear genetic programming approach. *Journal of Hydroinformatics*, 16(6), 1318–1330.
- Wang, W.-c., Xu, D.-m., Chau, K.-w., & Lei, G.-j. (2014). Assessment of river water quality based on theory of variable fuzzy sets and fuzzy binary comparison method. *Water Resources Management*, 28(12), 4183–4200.
- Wu, C. L., & Chau, K. W. (2006). A flood forecasting neural network model with genetic algorithm. *International Journal of Environment and Pollution*, 28(3–4), 261–273.
- Yang, J., Zeng, X., Zhong, S., & Wu, S. (2013). Effective neural network ensemble approach for improving generalization performance. *IEEE Transactions on Neural Networks and Learning Systems*, 24(6), 878–887.
- Zhai, J.-h., Xu, H.-y., & Wang, X.-z. (2012). Dynamic ensemble extreme learning machine based on sample entropy. *Soft Computing*, 16(9), 1493–1502.